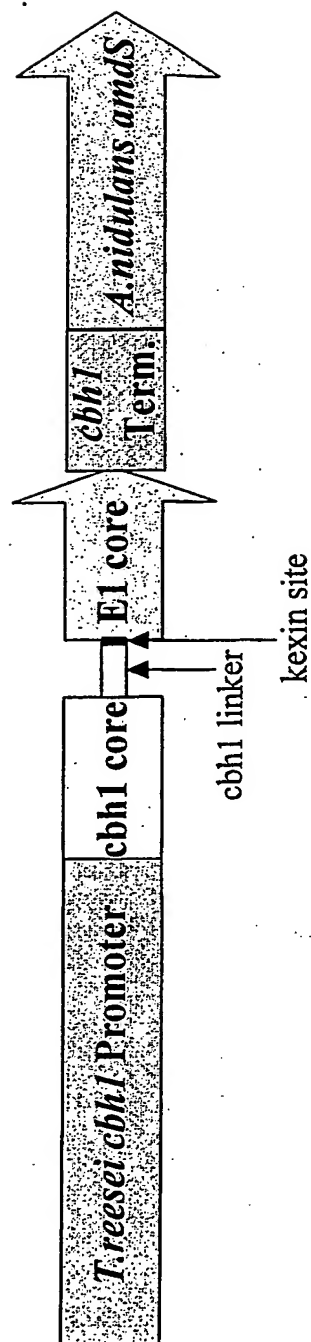FIGURE 1

CBH1-E1 Fusion Construct

T.reesei cbh1 core, linker (no CBD)

+

Acidothermus cellulolyticus endoglucanase 1 core (E1)

Figure 2

DNA sequence of *T.reesei cbh1* signal sequence+catalytic domain+linker (1570 bases)

ATGTATCGGAAGTTGGCCGTCATCTCGGCCTTCTTGGCCACAGCTCGTGCTCA
GTCGGCCTGCACTCTCCAATCGGAGACTCACCCGCCTCTGACATGGCAG
AAATGCTCGTCTGGTGGCACTTGCACTCAACAGACAGGCTCCGTGGTCA
TCGACGCCAACTGGCGCTGGACTCACGCTACGAACAGCAGCACGAACTG
CTACGATGGCAACACTTGGAGCTCGACCCTATGTCCTGACAACGAGACC
TGCGCGAAGAACTGCTGTCTGGACGGTGCCGCCTACGCGTCCACGTACG
GAGTTACCACGAGCGGTAACAGCCTCTCCATTGGCTTTGTCACCCAGTC
TGCGCAGAAGAACGTTGGCGCTCGCCTTTACCTATGGCGAGCGACACG
ACCTACCAGGAATTCACCCTGCTTGGCAACGAGTTCTCTTTCGATGTTGA
TGTTTCGCAGCTGCCGTAAGTGACTTACCATGAACCCCTGACGTATCTTC
TTGTGGGCTCCCAGCTGACTGGCCAATTTAAGGTGCGGCTTGAACGGAG
CTCTCTACTTCGTGTCCATGGACGCGGATGGTGGCGTGAGCAAGTATCC
CACCAACACCGCTGGCGCCAAGTACGGCACGGGGTACTGTGACAGCCAG
TGTCCCCGCGATCTGAAGTTCATCAATGGCCAGGCCAACGTTGAGGGCT
GGGAGCCGTCATCCAACAACGCAAACACGGGCATTGGAGGACACGGAA
GCTGCTGCTCTGAGATGGATATCTGGGAGGCCAACTCCATCTCCGAGGC
TCTTACCCCCCACCCTTGCACGACTGTCGGCCAGGAGATCTGCGAGGGT
GATGGGTGCGGCGGAACTTACTCCGATAACAGATATGGCGGCACTTGCG
ATCCCGATGGCTGCGACTGGAACCCATACCGCCTGGGCAACACCAGCTT
CTACGGCCCTGGCTCAAGCTTTACCCTCGATACCACCAAGAAATTGACC
GTTGTCACCCAGTTCGAGACGTCGGGTGCCATCAACCGATACTATGTCC
AGAATGGCGTCACTTTCCAGCAGCCCAACGCCGAGCTTGGTAGTTACTC
TGGCAACGAGCTCAACGATGATTACTGCACAGCTGAGGAGGCAGAATTC
GGCGGATCCTCTTTCTCAGACAAGGGCGGCCTGACTCAGTTCAAGAAGG
CTACCTCTGGCGGCATGGTTCTGGTCATGAGTCTGTGGGATGATGTGAG
TTTGATGGACAAACATGCGCGTTGACAAAGAGTCAAGCAGCTGACTGAG
ATGTTACAGTACTACGCCAACATGCTGTGGCTGGACTCCACCTACCCGA
CAAACGAGACCTCCTCCACACCCGGTGCCGTGCGCGGAAGCTGCTCCAC
CAGCTCCGGTGTCCCTGCTCAGGTCGAATCTCAGTCTCCCAACGCCAAG
GTCACCTTCTCCAACATCAAGTTCGGACCCATTGGCAGCACCGGCAACC
CTAGCGGCGGCAACCCTCCCGGCGGAAACCCGCCTGGCACCACCACCACCCG
CCGCCCAGCCACTACCACTGGAAGCTCTCCCGGACCTACTAGT

Figure 3

Amino acid sequence of *T.reesei cbh1* signal sequence + catalytic domain + linker (480 amino acids)

MYRKLAVISAFLATARAQSACTLQSETHPPLTWQKCSSGGTCTQQTGSVVID
ANWRWTHATNSSTNCYDGNTWSSTLCPDNETCAKNCCLDGAAYASTYGVT
TSGNSLSIGFVTQSAQKNVGARLYLMASDTTYQEFTLLGNEFSFDVDVSQLP
CGLNGALYFVSMDADGGVSKYPTNTAGAKYGTGYCDSQCPRDLKFINGQA
NVEGWEPSSNNANTGIGGHGSCCSEMDIWEANSISEALTPHPCTTVGQEICE
GDGCGGTYSDNRYGGTCDPDGCDWNPYRLGNTSFYGPGSSFTLDTTKKLT
VVTQFETSGAINRYYVQNGVTFQQPNAELGSYSGNELNDDYCTAEEAEFGG
SSFSDKGGLTQFKKATSGGMVLVMSLWDDYYANMLWLDSTYPTNETSSTP
GAVRGSCSTSSGVPAQVESQSPNAKVTFSNIKFGPIGSTGN*PSGGNPPGGNPPG*
TTTTRRPATTTGSSPGPTS

Figure 4

DNA sequence of *Acidothermus cellulolyticus GH5A* endoglucanase *1* catalytic domain
(1077 bases)

```
GCGGGCGGCGGCTATTGGCACACGAGCGGCCGGGAGATCCTGGACGCGAAC
AACGTGCCGGTACGGATCGCCGGCATCAACTGGTTTGGGTTCGAAACCTGCA
ATTACGTCGTGCACGGTCTCTGGTCACGCGACTACCGCAGCATGCTCGACCA
GATAAAGTCGCTCGGCTACAACACAATCCGGCTGCCGTACTCTGACGACATT
CTCAAGCCGGGCACCATGCCGAACAGCATCAATTTTTACCAGATGAATCAGG
ACCTGCAGGGTCTGACGTCCTTGCAGGTCATGGACAAAATCGTCGCGTACGC
CGGTCAGATCGGCCTGCGCATCATTCTTGACCGCCACCGACCGGATTGCAGC
GGGCAGTCGGCGCTGTGGTACACGAGCAGCGTCTCGGAGGCTACGTGGATTT
CCGACCTGCAAGCGCTGGCGCAGCGCTACAAGGGAAACCCGACGGTCGTCG
GCTTTGACTTGCACAACGAGCCGCATGACCCGGCCTGCTGGGGCTGCGGCGA
TCCGAGCATCGACTGGCGATTGGCCGCCGAGCGGGCCGGAAACGCCGTGCTC
TCGGTGAATCCGAACCTGCTCATTTTCGTCGAAGGTGTGCAGAGCTACAACG
GAGACTCCTACTGGTGGGGCGGCAACCTGCAAGGAGCCGGCCAGTACCCGGT
CGTGCTGAACGTGCCGAACCGCCTGGTGTACTCGGCGCACGACTACGCGACG
AGCGTCTACCCGCAGACGTGGTTCAGCGATCCGACCTTCCCCAACAACATGC
CCGGCATCTGGAACAAGAACTGGGGATACCTCTTCAATCAGAACATTGCACC
GGTATGGCTGGGCGAATTCGGTACGACACTGCAATCCACGACCGACCAGACG
TGGCTGAAGACGCTCGTCCAGTACCTACGGCCGACCGCGCAATACGGTGCGG
ACAGCTTCCAGTGGACCTTCTGGTCCTGGAACCCCGATTCCGGCGACACAGG
AGGAATTCTCAAGGATGACTGGCAGACGGTCGACACAGTAAAAGACGGCTAT
CTCGCGCCGATCAAGTCGTCGATTTTCGATCCTGTCGGC
```

Figure 5

Amino acid sequence of *Acidothermus cellulolyticus* GH5A endoglucanase 1 catalytic domain (359 amino acids)

AGGGYWHTSGREILDANNVPVRIAGINWFGFETCNYVVHGLWSRDYRSMLDQI
KSLGYNTIRLPYSDDILKPGTMPNSINFYQMNQDLQGLTSLQVMDKIVAYAGQIG
LRIILDRHRPDCSGQSALWYTSSVSEATWISDLQALAQRYKGNPTVVGFDLHNEP
HDPACWGCGDPSIDWRLAAERAGNAVLSVNPNLLIFVEGVQSYNGDSYWWGG
NLQGAGQYPVVLNVPNRLVYSAHDYATSVYPQTWFSDPTFPNNMPGIWNKNW
GYLFNQNIAPVWLGEFGTTLQSTTDQTWLKTLVQYLRPTAQYGADSFQWTFWS
WNPDSGDTGGILKDDWQTVDTVKDGYLAPIKSSIFDPVG

FIGURE 6

DNA sequence of *Acidothermus cellulolyticus GH48 cellulase* catalytic domain (1914 bases)

```
AACGACCCGTACATCCAGCGGTTCCTCACGATGTACAACAAGATTCACGACC
CAGCGAACGGCTACTTCAGCCCGCAGGGAATTCCCTACCACTCGGTAGAAAC
GCTCATCGTTGAGGCACCGGACTACGGGCACGAGACAACTTCGGAGGCGTAC
AGCTTCTGGCTCTGGCTCGAAGCGACGTACGGCGCAGTGACCGGCAACTGGA
CGCCGTTCAACAACGCCTGGACGACGATGGAAACGTACATGATCCCGCAGCA
CGCGGACCAGCCGAACAACGCGTCGTACAACCCCAACAGCCCGGCGTCGTAC
GCTCCGGAAGAGCCGCTGCCCAGCATGTACCCGGTTGCCATCGACAGCAGCG
TGCCGGTTGGGCACGACCCGCTCGCCGCCGAATTGCAGTCGACGTACGGCAC
TCCGGACATTTACGGCATGCACTGGCTGGCCGACGTTGACAACATCTACGGA
TACGGCGACAGCCCCGGCGGTGGTTGCGAACTCGGTCCTTCCGCTAAGGGCG
TCTCCTACATCAACACATTCCAGCGCGGCTCGCAGGAGTCCGTCTGGGAGAC
GGTCACCCAGCCGACGTGCGACAACGGCAAGTACGGTGGGGCGCACGGCTA
CGTCGACCTGTTCATCCAGGGTTCGACGCCGCCGCAGTGGAAGTACACCGAT
GCCCCGGACGCCGACGCCCGTGCCGTCCAGGCTGCGTACTGGGCCTACACCT
GGGCATCGGCGCAGGGCAAGGCAAGCGCGATTGCCCCGACGATCGCCAAGG
CGGCCAAACtCGGCGACTACCTGCGGTACTCGCTCTTTGACAAGTACTTCAAG
CAGGTCGGCAACTGCTACCCGGCCAGCTCCTGCCCTGGAGCAACCGGACGCC
AGAGCGAGACCTACCTGATCGGCTGGTACTACGCCTGGGGCGGCTCAAGCCA
AGGCTGGGCCTGGCGCATTGGTGACGGCGCCGCGCACTTCGGCTACCAGAAT
CCGCTTGCCGCGTGGGCGATGTCGAACGTGACACCGCTCATTCCGCTCTCGCC
CACGGCAAAGAGCGACTGGGCGGCGAGCTTGCAGCGCCAGCTGGAGTTCTAC
CAGTGGTTGCAATCCGCGGAAGGAGCCATTGCGGGCGGCGCCACCAACAGCT
GGAACGGCAATTACGGGACCCCGCCGGCCGGAGACTCGACCTTCTACGGCAT
GGCGTACGACTGGGAGCCGGTCTACCACGACCCGCCGAGCAACAACTGGTTC
GGCTTCCAGGCGTGGTCCATGGAACGGGTTGCCGAGTACTACTACGTCACCG
GCGACCCGAAGGCCAAGGCGCTGCTCGACAAGTGGGTCGCATGGGTGAAGC
CGAATGTCACCACCGGTGCCTCATGGTCGATTCCGTCGAATTTGTCCTGGAGC
GGCCAACCGGATACCTGGAATCCGAGCAACCCAGGAACGAATGCCAACCTG
CACGTGACCATCACGTCGTCCGGGCAGGACGTCGGTGTTGCCGCGGCGCTCG
CGAAGACACTCGAGTACTACGCGGCAAAATCCGGCGATACGGCCTCGCGCGA
CCTCGCGAAGGGATTGCTCGACTCCATGTGGAACAACGACCAGGACAGCCTC
GGTGTGAGCACACCGGAGACGCGGACCGACTACTCTCGGTTCACTCAGGTGT
ACGACCCGACGACTGGTGACGGCCTCTACATCCCGTCGGGTTGGACGGGGAC
CATGCCCAACGGTGACCAAATCAAGCCGGGTGCGACCTTCCTGAGCATCCGG
TCCTGGTACACCAAGGATCCGCAGTGGTCGAAGGTGCAGGCGTACCTCAACG
GCGGGCCTGCTCCGACGTTCAACTACCACCGGTTCTGGGCGGAGTCCGACTT
CGCGATGGCGAACGCCGATTTTGGCATGCTCTTCCCATCCGGG
```

FIGURE 7

Amino acid sequence of *Acidothermus cellulolyticus 48* catalytic domain (638 amino acids)

NDPYIQRFLTMYNKIHDPANGYFSPQGIPYHSVETLIVEAPDYGHETTSEAYSFW
LWLEATYGAVTGNWTPFNNAWTTMETYMIPQHADQPNNASYNPNSPASYAPEE
PLPSMYPVAIDSSVPVGHDPLAAELQSTYGTPDIYGMHWLADVDNIYGYGDSPG
GGCELGPSAKGVSYINTFQRGSQESVWETVTQPTCDNGKYGGAHGYVDLFIQGS
TPPQWKYTDAPDADARAVQAAYWAYTWASAQGKASAIAPTIAKAAKLGDYLR
YSLFDKYFKQVGNCYPASSCPGATGRQSETYLIGWYYAWGGSSQGWAWRIGD
GAAHFGYQNPLAAWAMSNVTPLIPLSPTAKSDWAASLQRQLEFYQWLQSAEGA
IAGGATNSWNGNYGTPPAGDSTFYGMAYDWEPVYHDPPSNNWFGFQAWSMER
VAEYYYVTGDPKAKALLDKWVAWVKPNVTTGASWSIPSNLSWSGQPDTWNPS
NPGTNANLHVTITSSGQDVGVAAALAKTLEYYAAKSGDTASRDLAKGLLDSMW.
NNDQDSLGVSTPETRTDYSRFTQVYDPTTGDGLYIPSGWTGTMPNGDQIKPGAT
FLSIRSWYTKDPQWSKVQAYLNGGPAPTFNYHRFWAESDFAMANADFGMLFPS
G

2

FIGURE 8A

DNA sequence of *Acidothermus cellulolyticus GH74* catalytic domain

GCGACGACTCAGCCGTACACCTGGAGCAACGTGGCGATCGGGGGGCGGCGGC
TTTGTCGACGGGATCGTCTTCAATGAAGGTGCACCGGGAATTCTGTACGTGCG
GACGGACATCGGGGGGGATGTATCGATGGGATGCCGCCAACGGGCGGTGGAT
CCCTCTTCTGGATTGGGTGGGATGGAACAATTGGGGGTACAACGGCGTCGTC
AGCATTGCGGCAGACCCGATCAATACTAACAAGGTATGGGCCGCCGTCGGAA
TGTACACCAACAGCTGGGACCCAAACGACGGAGCGATTCTCCGCTCGTCTGA
TCAGGGCGCAACGTGGCAAATAACGCCCCTGCCGTTCAAGCTTGGCGGCAAC
ATGCCCGGGCGTGGAATGGGCGAGCGGCTTGCGGTGGATCCAAACAATGACA
ACATTCTGTATTTCGGCGCCCCGAGCGGCAAAGGGCTCTGGAGAAGCACAGA
TTCCGGCGCGACCTGGTCCCAGATGACGAACTTTCCGGACGTAGGCACGTAC
ATTGCAAATCCCACTGACACGACCGGCTATCAGAGCGATATTCAAGGCGTCG
TCTGGGTCGCTTTCGACAAGTCTTCGTCATCGCTCGGGCAAGCGAGTAAGACC
ATTTTTGTGGGCGTGGCGGATCCCAATAATCCGGTCTTCTGGAGCAGAGACG
GCGGCGCGACGTGGCAGGCGGTGCCGGGTGCGCCGACCGGCTTCATCCCGCA
CAAGGGCGTCTTTGACCCGGTCAACCACGTGCTCTATATTGCCACCAGCAAT
ACGGGTGGTCCGTATGACGGGAGCTCCGGCGACGTCTGGAAATTCTCGGTGA
CCTCCGGGACATGGACGCGAATCAGCCCGGTACCTTCGACGGACACGGCCAA
CGACTACTTTGGTTACAGCGGCCTCACTATCGACCGCCAGCACCCGAACACG
ATAATGGTGGCAACCCAGATATCGTGGTGGCCGGACACCATAATCTTTCGGA
GCACCGACGGCGGTGCGACGTGGACGCGGATCTGGGATTGGACGAGTTATCC
CAATCGAAGCTTGCGATATGTGCTTGACATTCGGCGGAGCCTTGGCTGACCT
TCGGCGTACAGCCGAATCCTCCCGTACCGAGTCCGAAGCTCGGCTGGATGGA
TGAAGCGATGGCAATCGATCCGTTCAACTCTGATCGGATGCTCTACGGAACA
GGCGCGACGTTGTACGCAACAAATGATCTCACGAAGTGGGACTCCGGCGGCC
AGATTCATATCGCGCCGATGGTCAAAGGATTGGAGGAGACGGCGGTAAACG
ATCTCATCAGCCCGCCGTCTGGCGCCCCGCTCATCAGCGCTCTCGGAGACCTC
GGCGGCTTCACCCACGCCGACGTTACTGCCGTGCCATCGACGATCTTCACGTC

FIGURE 8B

ACCGGTGTTCACGACCGGCACCAGCGTCGACTATGCGGAATTGAATCCGTCG
ATCATCGTTCGCGCTGGAAGTTTCGATCCATCGAGCCAACCGAACGACAGGC
ACGTCGCGTTCTCGACAGACGGCGGCAAGAACTGGTTCCAAGGCAGCGAACC
TGGCGGGGTGACGACGGGCGGCACCGTCGCCGCATCGGCCGACGGCTCTCGT
TTCGTCTGGGCTCCCGGCGATCCCGGTCAGCCTGTGGTGTACGCAGTCGGATT
TGGCAACTCCTGGGCTGCTTCGCAAGGTGTTCCCGCCAATGCCCAGATCCGCT
CAGACCGGGTGAATCCAAAGACTTTCTATGCCCTATCCAATGGAACCTTCTAT
CGAAGCACGGACGGCGGCGTGACATTCCAACCGGTCGCGGCCGGTCTTCCGA
GCAGCGGTGCCGTCGGTGTCATGTTCCACGCGGTGCCTGGAAAAGAAGGCGA
TCTGTGGCTCGCTGCATCGAGCGGGCTTTACCACTCAACCAATGGCGGCAGC
AGTTGGTCTGCAATCACCGGCGTATCCTCCGCGGTGAACGTGGGATTTGGTA
AGTCTGCGCCCGGGTCGTCATACCCAGCCGTCTTTGTCGTCGGCACGATCGGA
GGCGTTACGGGGGCGTACCGCTCCGACGACGGTGGGACGACCTGGGTACGG
ATCAATGATGACCAGCACCAATACGGAAATTGGGGACAAGCAATCACCGGTG
ACCCGCGAATTTACGGGCGGGTGTACATAGGCACGAACGGCCGTGGAATTGT
CTACGGGGACATTGGTGGTGCGCCGTCCGGATCG

FIGURE 9

Amino acid sequence of *Acidothermus cellulolyticus 74* catalytic domain (741 amino acids)

ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPL
LDWVGWNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQ
GATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSG
ATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVG
VADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPY
DGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQIS
WWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPS
PKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLE
ETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNP
SIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRS
TDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWS
AITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDGGTTWVRINDDQ
HQYGNWGQAITGDPRIYGRVYIGTNGRGIVYGDIGGAPSGS

Figure 10

DNA sequence of *Thermobifida fusca* E3 (TfE3) cellulase including the cellulose binding domain - linker – cataltyic domain but lacking a TfE3 signal sequence. (1677 bases)

```
GCCGGCTGCTCGGTGGACTACACGGTCAACTCCTGGGGTACCGGGTTCACCG
CCAACGTCACCATCACCAACCTCGGCAGTGCGATCAACGGCTGGACCCTGGA
GTGGGACTTCCCCGGCAACCAGCAGGTGACCAACCTGTGGAACGGGACCTAC
ACCCAGTCCGGGCAGCACGTGTCGGTCAGCAACGCCCCGTACAACGCCTCCA
TCCCGGCCAACGGAACGGTTGAGTTCGGGTTCAACGGCTCCTACTCGGGCAG
CAACGACATCCCCTCCTCCTTCAAGCTGAACGGGGTTACCTGCGACGGCTCG
GACGACCCCGACCCCGAGCCCAGCCCCTCCCCCAGCCCTTCCCCCAGCCCCA
CAGACCCGGATGAGCCGGGCGGCCCGACCAACCCGCCCACCAACCCCGGCG
AGAAGGTCGACAACCCGTTCGAGGGCGCCAAGCTGTACGTGAACCCGGTCTG
GTCGGCCAAGGCCGCCGCTGAGCCGGGCGGTTCCGCGGTCGCCAACGAGTCC
ACCGCTGTCTGGCTGGACCGTATCGGCGCCATCGAGGGCAACGACAGCCCGA
CCACCGGCTCCATGGGTCTGCGCGACCACCTGGAGGAGGCCGTCCGCCAGTC
CGGTGGCGACCCGCTGACCATCCAGGTCGTCATCTACAACCTGCCCGGCCGC
GACTGCGCCGCGCTGGCCTCCAACGGTGAGCTGGGTCCCGATGAACTCGACC
GCTACAAGAGCGAGTACATCGACCCGATCGCCGACATCATGTGGGACTTCGC
AGACTACGAGAACCTGCGGATCGTCGCCATCATCGAGATCGACTCCCTGCCC
AACCTCGTCACCAACGTGGGCGGGAACGGCGGCACCGAGCTCTGCGCCTACA
TGAAGCAGAACGGCGGCTACGTCAACGGTGTCGGCTACGCCCTCCGCAAGCT
GGGCGAGATCCCGAACGTCTACAACTACATCGACGCCGCCCACCACGGCTGG
ATCGGCTGGGACTCCAACTTCGGCCCCTCGGTGGACATCTTCTACGAGGCCG
CCAACGCCTCCGGCTCCACCGTGGACTACGTGCACGGCTTCATCTCCAACAC
GGCCAACTACTCGGCCACTGTGGAGCCGTACCTGGACGTCAACGGCACCGTT
AACGGCCAGCTCATCCGCCAGTCCAAGTGGGTTGACTGGAACCAGTACGTCG
ACGAGCTCTCCTTCGTCCAGGACCTGCGTCAGGCCCTGATCGCCAAGGGCTTC
CGGTCCGACATCGGTATGCTCATCGACACCTCCCGCAACGGCTGGGGTGGCC
CGAACCGTCCGACCGGACCGAGCTCCTCCACCGACCTCAACACCTACGTTGA
CGAGAGCCGTATCGACCGCCGTATCCACCCCGGTAACTGGTGCAACCAGGCC
GGTGCGGGCCTCGGCGAGCGGCCCACGGTCAACCCGGCTCCCGGTGTTGACG
CCTACGTCTGGGTGAAGCCCCGGGTGAGTCCGACGGCGCCAGCGAGGAGAT
CCCGAACGACGAGGGCAAGGGCTTCGACCGCATGTGCGACCCGACCTACCAG
GGCAACGCCCGCAACGGCAACAACCCCTCGGGTGCGCTGCCCAACGCCCCCA
TCTCCGGCCACTGGTTCTCTGCCCAGTTCCGCGAGCTGCTGGCCAACGCCTAC
CCGCCTCTGTAA
```

Figure 11

Amino acid sequence of the *Thermobifida fusca E3* - cellulase including the cellulose binding domain - linker – cataltyic domain but lacking the TfE3 signal sequence. (558 amino acids)

AGCSVDYTVNSWGTGFTANVTITNLGSAINGWTLEWDFPGNQQVTNLWNGTYT
QSGQHVSVSNAPYNASIPANGTVEFGFNGSYSGSNDIPSSFKLNGVTCDGSDDPD
PEPSPSPSPSPSPTDPDEPGGPTNPPTNPGEKVDNPFEGAKLYVNPVWSAKAAAEP
GGSAVANESTAVWLDRIGAIEGNDSPTTGSMGLRDHLEEAVRQSGGDPLTIQVVI
YNLPGRDCAALASNGELGPDELDRYKSEYIDPIADIMWDFADYENLRIVAIIEIDS
LPNLVTNVGGNGGTELCAYMKQNGGYVNGVGYALRKLGEIPNVYNYIDAAHH
GWIGWDSNFGPSVDIFYEAANASGSTVDYVHGFISNTANYSATVEPYLDVNGTV
NGQLIRQSKWVDWNQYVDELSFVQDLRQALIAKGFRSDIGMLIDTSRNGWGGP
NRPTGPSSSTDLNTYVDESRIDRRIHPGNWCNQAGAGLGERPTVNPAPGVDAYV
WVKPPGESDGASEEIPNDEGKGFDRMCDPTYQGNARNGNNPSGALPNAPISGH
WFSAQFRELLANAYPPL

Figure 12

DNA sequence of *Thermobifida fusca E5* (TfE5) endoglucanase including the cellulose binding domain - linker and catalytic domain but lacking a TfE5 signal sequence. (1293 bases)

```
GCCGGTCTCACCGCCACAGTCACCAAAGAATCCTCGTGGGACAACGGCTACT
CCGCGTCCGTCACCGTCCGCAACGACACCTCGAGCACCGTCTCCCAGTGGGA
GGTCGTCCTCACCCTGCCCGGCGGCACTACAGTGGCCCAGGTGTGGAACGCC
CAGCACACCAGCAGCGGCAACTCCCACACCTTCACCGGGGTTTCCTGGAACA
GCACCATCCCGCCCGGAGGCACCGCCTCTTCCGGCTTCATCGCTTCCGGCAGC
GGCGAACCCACCCACTGCACCATCAACGGCGCCCCCTGCGACGAAGGCTCCG
AGCCGGGCGGCCCCGGCGGTCCCGGAACCCCCTCCCCCGACCCCGGCACGCA
GCCCGGCACCGGCACCCCGGTCGAGCGGTACGGCAAAGTCCAGGTCTGCGGC
ACCCAGCTCTGCGACGAGCACGGCAACCCGGTCCAACTGCGCGGCATGAGCA
CCCACGGCATCCAGTGGTTCGACCACTGCCTGACCGACAGCTCGCTGGACGC
CCTGGCCTACGACTGGAAGGCCGACATCATCCGCCTGTCCATGTACATCCAG
GAAGACGGCTACGAGACCAACCCGCGCGGCTTCACCGACCGGATGCACCAG
CTCATCGACATGGCCACGGCGCGCGGCCTGTACGTGATCGTGGACTGGCACA
TCCTCACCCCGGGCGATCCCCACTACAACCTGGACCGGGCCAAGACCTTCTTC
GCGGAAATCGCCCAGCGCCACGCCAGCAAGACCAACGTGCTCTACGAGATCG
CCAACGAACCCAACGGAGTGAGCTGGGCCTCCATCAAGAGCTACGCCGAAG
AGGTCATCCCGGTGATCCGCCAGCGCGACCCCGACTCGGTGATCATCGTGGG
CACCCGCGGCTGGTCGTCGCTCGGCGTCTCCGAAGGCTCCGGCCCCGCCGAG
ATCGCGGCCAACCCGGTCAACGCCTCCAACATCATGTACGCCTTCCACTTCTA
CGCGGCCTCGCACCGCGACAACTACCTCAACGCGCTGCGTGAGGCCTCCGAG
CTGTTCCCGGTCTTCGTCACCGAGTTCGGCACCGAGACCTACACCGGTGACG
GCGCCAACGACTTCCAGATGGCCGACCGCTACATCGACCTGATGGCGGAACG
GAAGATCGGGTGGACCAAGTGGAACTACTCGGACGACTTCCGTTCCGGCGCG
GTCTTCCAGCCGGGCACCTGCGCGTCCGGCGGCCCGTGGAGCGGTTCGTCGC
TGAAGGCGTCCGGACAGTGGGTGCGGAGCAAGCTCCAGTCCTGA
```

Figure 13

Amino acid sequence of the *Thermobifida fusca E5* –cellulase including the cellulose binding domain - linker - catalytic domain but lacking a TfE5 signal sequence. (430 amino acids)

AGLTATVTKESSWDNGYSASVTVRNDTSSTVSQWEVVLTLPGGTTVAQVWNAQ
HTSSGNSHTFTGVSWNSTIPPGGTASSGFIASGSGEPTHCTINGAPCDEGSEPGGP
GGPGTPSPDPGTQPGTGTPVERYGKVQVCGTQLCDEHGNPVQLRGMSTHGIQW
FDHCLTDSSLDALAYDWKADIIRLSMYIQEDGYETNPRGFTDRMHQLIDMATAR
GLYVIVDWHILTPGDPHYNLDRAKTFFAEIAQRHASKTNVLYEIANEPNGVSWA
SIKSYAEEVIPVIRQRDPDSVIIVGTRGWSSLGVSEGSGPAEIAANPVNASNIMYAF
HFYAASHRDNYLNALREASELFPVFVTEFGTETYTGDGANDFQMADRYIDLMA
ERKIGWTKWNYSDDFRSGAVFQPGTCASGGPWSGSSLKASGQWVRSKLQS

Figure 14

DNA sequence of CBH1-E1 fusion (2656 bases)
T.reesei cbh1 signal sequence+catalytic domain+linker+added amino acids
SKR+*Acidothermus cellulolyticus* GH5A catalytic domain

```
ATGTATCGGAAGTTGGCCGTCATCTCGGCCTTCTTGGCCACAGCTCGTGCTCA
GTCGGCCTGCACTCTCCAATCGGAGACTCACCCGCCTCTGACATGGCAGAAA
TGCTCGTCTGGTGGCACTTGCACTCAACAGACAGGCTCCGTGGTCATCGACG
CCAACTGGCGCTGGACTCACGCTACGAACAGCAGCACGAACTGCTACGATGG
CAACACTTGGAGCTCGACCCTATGTCCTGACAACGAGACCTGCGCGAAGAAC
TGCTGTCTGGACGGTGCCGCCTACGCGTCCACGTACGGAGTTACCACGAGCG
GTAACAGCCTCTCCATTGGCTTTGTCACCCAGTCTGCGCAGAAGAACGTTGGC
GCTCGCCTTTACCTTATGGCGAGCGACACGACCTACCAGGAATTCACCCTGCT
TGGCAACGAGTTCTCTTTCGATGTTGATGTTTCGCAGCTGCCGTAAGTGACTT
ACCATGAACCCCTGACGTATCTTCTTGTGGGCTCCCAGCTGACTGGCCAATTT
AAGGTGCGGCTTGAACGGAGCTCTCTACTTCGTGTCCATGGACGCGGATGGT
GGCGTGAGCAAGTATCCCACCAACACCGCTGGCGCCAAGTACGGCACGGGGT
ACTGTGACAGCCAGTGTCCCCGCGATCTGAAGTTCATCAATGGCCAGGCCAA
CGTTGAGGGCTGGGAGCCGTCATCCAACAACGCAAACACGGGCATTGGAGG
ACACGGAAGCTGCTGCTCTGAGATGGATATCTGGGAGGCCAACTCCATCTCC
GAGGCTCTTACCCCCCACCCTTGCACGACTGTCGGCCAGGAGATCTGCGAGG
GTGATGGGTGCGGCGGAACTTACTCCGATAACAGATATGGCGGCACTTGCGA
TCCCGATGGCTGCGACTGGAACCCATACCGCCTGGGCAACACCAGCTTCTAC
GGCCCTGGCTCAAGCTTTACCCTCGATACCACCAAGAAATTGACCGTTGTCAC
CCAGTTCGAGACGTCGGGTGCCATCAACCGATACTATGTCCAGAATGGCGTC
ACTTTCCAGCAGCCCAACGCCGAGCTTGGTAGTTACTCTGGCAACGAGCTCA
ACGATGATTACTGCACAGCTGAGGAGGCAGAATTCGGCGGATCCTCTTTCTC
AGACAAGGGCGGCCTGACTCAGTTCAAGAAGGCTACCTCTGGCGGCATGGTT
CTGGTCATGAGTCTGTGGGATGATGTGAGTTTGATGGACAAACATGCGCGTT
GACAAAGAGTCAAGCAGCTGACTGAGATGTTACAGTACTACGCCAACATGCT
GTGGCTGGACTCCACCTACCCGACAAACGAGACCTCCTCCACACCCGGTGCC
GTGCGCGGAAGCTGCTCCACCAGCTCCGGTGTCCCTGCTCAGGTCGAATCTC
AGTCTCCCAACGCCAAGGTCACCTTCTCCAACATCAAGTTCGGACCCATTGGC
AGCACCGGCAACCCTAGCGGCGGCAACCCTCCCGGCGGAAACCCGCCTGGCA
CCACCACCACCCGCCGCCCAGCCACTACCACTGGAAGCTCTCCCGGACCTAC
TAGTAAGCGGGCGGGCGGCGGCTATTGGCACACGAGCGGCCGGGAGATCCT
GGACGCGAACAACGTGCCGGTACGGATCGCCGGCATCAACTGGTTTGGGTTC
GAAACCTGCAATTACGTCGTGCACGGTCTCTGGTCACGCGACTACCGCAGCA
TGCTCGACCAGATAAAGTCGCTCGGCTACAACACAATCCGGCTGCCGTACTC
TGACGACATTCTCAAGCCGGGCACCATGCCGAACAGCATCAATTTTACCAG
ATGAATCAGGACCTGCAGGGTCTGACGTCCTTGCAGGTCATGGACAAAATCG
TCGCGTACGCCGGTCAGATCGGCCTGCGCATCATTCTTGACCGCCACCGACC
GGATTGCAGCGGGCAGTCGGCGCTGTGGTACACGAGCAGCGTCTCGGAGGCT
ACGTGGATTTCCGACCTGCAAGCGCTGGCGCAGCGCTACAAGGGAAACCCGA
```

CGGTCGTCGGCTTTGACTTGCACAACGAGCCGCATGACCCGGCCTGCTGGGG
CTGCGGCGATCCGAGCATCGACTGGCGATTGGCCGCCGAGCGGGCCGGAAAC
GCCGTGCTCTCGGTGAATCCGAACCTGCTCATTTTCGTCGAAGGTGTGCAGAG
CTACAACGGAGACTCCTACTGGTGGGGCGGCAACCTGCAAGGAGCCGGCCA
GTACCCGGTCGTGCTGAACGTGCCGAACCGCCTGGTGTACTCGGCGCACGAC
TACGCGACGAGCGTCTACCCGCAGACGTGGTTCAGCGATCCGACCTTCCCCA
ACAACATGCCCGGCATCTGGAACAAGAACTGGGGATACCTCTTCAATCAGAA
CATTGCACCGGTATGGCTGGGCGAATTCGGTACGACACTGCAATCCACGACC
GACCAGACGTGGCTGAAGACGCTCGTCCAGTACCTACGGCCGACCGCGCAAT
ACGGTGCGGACAGCTTCCAGTGGACCTTCTGGTCCTGGAACCCCGATTCCGG
CGACACAGGAGGAATTCTCAAGGATGACTGGCAGACGGTCGACACAGTAAA
AGACGGCTATCTCGCGCCGATCAAGTCGTCGATTTTCGATCCTGTCGGCTAA

Figure 15

Amino acid sequence of CBH1-E1 fusion (841 amino acids)
T.reesei cbh1 signal sequence+catalytic domain+linker+added amino acids
SKR+*Acidothermus cellulolyticus* GH5A catalytic domain

MYRKLAVISAFLATARAQSACTLQSETHPPLTWQKCSSGGTCTQQTGSVVIDAN
WRWTHATNSSTNCYDGNTWSSTLCPDNETCAKNCCLDGAAYASTYGVTTSGNS
LSIGFVTQSAQKNVGARLYLMASDTTYQEFTLLGNEFSFDVDVSQLPCGLNGAL
YFVSMDADGGVSKYPTNTAGAKYGTGYCDSQCPRDLKFINGQANVEGWEPSSN
NANTGIGGHGSCCSEMDIWEANSISEALTPHPCTTVGQEICEGDGCGGTYSDNRY
GGTCDPDGCDWNPYRLGNTSFYGPGSSFTLDTTKKLTVVTQFETSGAINRYYVQ
NGVTFQQPNAELGSYSGNELNDDYCTAEEAEFGGSSFSDKGGLTQFKKATSGGM
VLVMSLWDDYYANMLWLDSTYPTNETSSTPGAVRGSCSTSSGVPAQVESQSPN
AKVTFSNIKFGPIGSTGNPSGGNPPGGNPPGTTTTRRPATTTGSSPGPTSKRAGGG
YWHTSGREILDANNVPVRIAGINWFGFETCNYVVHGLWSRDYRSMLDQIKSLGY
NTIRLPYSDDILKPGTMPNSINFYQMNQDLQGLTSLQVMDKIVAYAGQIGLRIILD
RHRPDCSGQSALWYTSSVSEATWISDLQALAQRYKGNPTVVGFDLHNEPHDPAC
WGCGDPSIDWRLAAERAGNAVLSVNPNLLIFVEGVQSYNGDSYWWGGNLQGA
GQYPVVLNVPNRLVYSAHDYATSVYPQTWFSDPTFPNNMPGIWNKNWGYLFN
QNIAPVWLGEFGTTLQSTTDQTWLKTLVQYLRPTAQYGADSFQWTFWSWNPDS
GDTGGILKDDWQTVDTVKDGYLAPIKSSIFDPVG

Figure 16
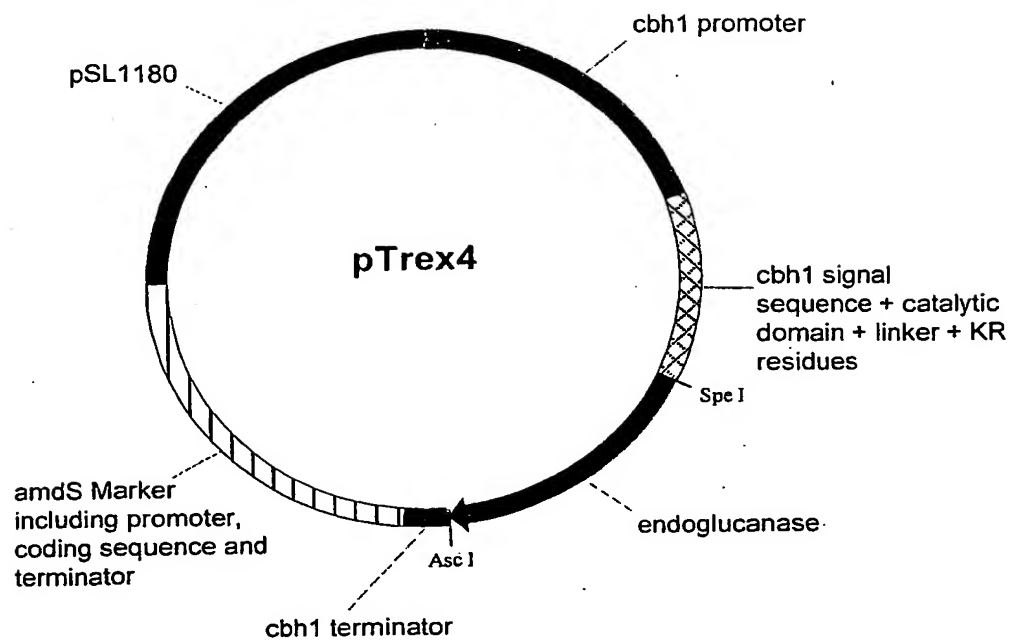
Figure 17

DNA sequence of pTrex4 (10239 bases)

AAGCTTAACTAGTACTTCTCGAGCTCTGTACATGTCCGGTCGCGACGTACGCG
TATCGATGGCGCCAGCTGCAGGCGGCCGCCTGCAGCCACTTGCAGTCCCGTG
GAATTCTCACGGTGAATGTAGGCCTTTTGTAGGGTAGGAATTGTCACTCAAGC
ACCCCCAACCTCCATTACGCCTCCCCCATAGAGTTCCCAATCAGTGAGTCATG
GCACTGTTCTCAAATAGATTGGGGAGAAGTTGACTTCCGCCCAGAGCTGAAG
GTCGCACAACCGCATGATATAGGGTCGGCAACGGCAAAAAGCACGTGGCT
CACCGAAAAGCAAGATGTTTGCGATCTAACATCCAGGAACCTGGATACATCC
ATCATCACGCACGACCACTTTGATCTGCTGGTAAACTCGTATTCGCCCTAAAC
CGAAGTGACGTGGTAAATCTACACGTGGGCCCCTTTCGGTATACTGCGTGTGT
CTTCTCTAGGTGCCATTCTTTTCCCTTCCTCTAGTGTTGAATTGTTTGTTGG
AGTCCGAGCTGTAACTACCTCTGAATCTCTGGAGAATGGTGGACTAACGACT
ACCGTGCACCTGCATCATGTATATAATAGTGATCCTGAGAAGGGGGGTTTGG
AGCAATGTGGGACTTTGATGGTCATCAAACAAAGAACGAAGACGCCTCTTTT
GCAAAGTTTTGTTTCGGCTACGGTGAAGAACTGGATACTTGTTGTGTCTTCTG
TGTATTTTTGTGGCAACAAGAGGCCAGAGACAATCTATTCAAACACCAAGCT
TGCTCTTTTGAGCTACAAGAACCTGTGGGGTATATATCTAGAGTTGTGAAGTC
GGTAATCCCGCTGTATAGTAATACGAGTCGCATCTAAATACTCCGAAGCTGCT
GCGAACCCGGAGAATCGAGATGTGCTGGAAAGCTTCTAGCGAGCGGCTAAAT
TAGCATGAAAGGCTATGAGAAATTCTGGAGACGGCTTGTTGAATCATGGCGT
TCCATTCTTCGACAAGCAAAGCGTTCCGTCGCAGTAGCAGGCACTCATTCCCG
AAAAAACTCGGAGATTCCTAAGTAGCGATGGAACCGGAATAATATAATAGGC
AATACATTGAGTTGCCTCGACGGTTGCAATGCAGGGGTACTGAGCTTGGACA
TAACTGTTCCGTACCCCACCTCTTCTCAACCTTTGGCGTTTCCCTGATTCAGCG
TACCCGTACAAGTCGTAATCACTATTAACCCAGACTGACCGGACGTGTTTTGC
CCTTCATTTGGAGAAATAATGTCATTGCGATGTGTAATTTGCCTGCTTGACCG
ACTGGGGCTGTTCGAAGCCCGAATGTAGGATTGTTATCCGAACTCTGCTCGTA
GAGGCATGTTGTGAATCTGTGTCGGGCAGGACACGCCTCGAAGGTTCACGGC

AAGGGAAACCACCGATAGCAGTGTCTAGTAGCAACCTGTAAAGCCGCAATGC
AGCATCACTGGAAAATACAAACCAATGGCTAAAAGTACATAAGTTAATGCCT
AAAGAAGTCATATACCAGCGGCTAATAATTGTACAATCAAGTGGCTAAACGT
ACCGTAATTTGCCAACGGCTTGTGGGGTTGCAGAAGCAACGGCAAAGCCCCA
CTTCCCCACGTTTGTTTCTTCACTCAGTCCAATCTCAGCTGGTGATCCCCAAT
TGGGTCGCTTGTTTGTTCCGGTGAAGTGAAAGAAGACAGAGGTAAGAATGTC
TGACTCGGAGCGTTTTGCATACAACCAAGGGCAGTGATGGAAGACAGTGAAA
TGTTGACATTCAAGGAGTATTTAGCCAGGGATGCTTGAGTGTATCGTGTAAG
GAGGTTTGTCTGCCGATACGACGAATACTGTATAGTCACTTCTGATGAAGTGG
TCCATATTGAAATGTAAGTCGGCACTGAACAGGCAAAGATTGAGTTGAAAC
TGCCTAAGATCTCGGGCCCTCGGGCCTTCGGCCTTTGGGTGTACATGTTTGTG
CTCCGGGCAAATGCAAAGTGTGGTAGGATCGAACACACTGCTGCCTTTACCA
AGCAGCTGAGGGTATGTGATAGGCAAATGTTCAGGGGCCACTGCATGGTTTC
GAATAGAAAGAGAAGCTTAGCCAAGAACAATAGCCGATAAAGATAGCCTCA
TTAAACGGAATGAGCTAGTAGGCAAAGTCAGCGAATGTGTATATAAAGGT
TCGAGGTCCGTGCCTCCCTCATGCTCTCCCCATCTACTCATCAACTCAGATCC
TCCAGGAGACTTGTACACCATCTTTTGAGGCACAGAAACCCAATAGTCAACC
GCGGACTGCGCATCATGTATCGGAAGTTGGCCGTCATCTCGGCCTTCTTGGCC
ACAGCTCGTGCTCAGTCGGCCTGCACTCTCCAATCGGAGACTCACCCGCCTCT
GACATGGCAGAAATGCTCGTCTGGTGGCACTTGCACTCAACAGACAGGCTCC
GTGGTCATCGACGCCAACTGGCGCTGGACTCACGCTACGAACAGCAGCACGA
ACTGCTACGATGGCAACACTTGGAGCTCGACCCTATGTCCTGACAACGAGAC
CTGCGCGAAGAACTGCTGTCTGGACGGTGCCGCCTACGCGTCCACGTACGGA
GTTACCACGAGCGGTAACAGCCTCTCCATTGGCTTTGTCACCCAGTCTGCGCA
GAAGAACGTTGGCGCTCGCCTTTACCTTATGGCGAGCGACACGACCTACCAG
GAATTCACCCTGCTTGGCAACGAGTTCTCTTTCGATGTTGATGTTTCGCAGCT
GCCGTAAGTGACTTACCATGAACCCCTGACGTATCTTCTTGTGGGCTCCCAGC
TGACTGGCCAATTTAAGGTGCGGCTTGAACGGAGCTCTCTACTTCGTGTCCAT
GGACGCGGATGGTGGCGTGAGCAAGTATCCCACCAACACCGCTGGCGCCAA
GTACGGCACGGGGTACTGTGACAGCCAGTGTCCCCGCGATCTGAAGTTCATC
AATGGCCAGGCCAACGTTGAGGGCTGGGAGCCGTCATCCAACAACGCAAAC

ACGGGCATTGGAGGACACGGAAGCTGCTGCTCTGAGATGGATATCTGGGAGG
CCAACTCCATCTCCGAGGCTCTTACCCCCCACCCTTGCACGACTGTCGGCCAG
GAGATCTGCGAGGGTGATGGGTGCGGCGGAACTTACTCCGATAACAGATATG
GCGGCACTTGCGATCCCGATGGCTGCGACTGGAACCCATACCGCCTGGGCAA
CACCAGCTTCTACGGCCCTGGCTCAAGCTTTACCCTCGATACCACCAAGAAAT
TGACCGTTGTCACCCAGTTCGAGACGTCGGGTGCCATCAACCGATACTATGTC
CAGAATGGCGTCACTTTCCAGCAGCCCAACGCCGAGCTTGGTAGTTACTCTG
GCAACGAGCTCAACGATGATTACTGCACAGCTGAGGAGGCAGAATTCGGCGG
ATCCTCTTTCTCAGACAAGGGCGGCCTGACTCAGTTCAAGAAGGCTACCTCTG
GCGGCATGGTTCTGGTCATGAGTCTGTGGGATGATGTGAGTTTGATGGACAA
ACATGCGCGTTGACAAAGAGTCAAGCAGCTGACTGAGATGTTACAGTACTAC
GCCAACATGCTGTGGCTGGACTCCACCTACCCGACAAACGAGACCTCCTCCA
CACCCGGTGCCGTGCGCGGAAGCTGCTCCACCAGCTCCGGTGTCCCTGCTCA
GGTCGAATCTCAGTCTCCCAACGCCAAGGTCACCTTCTCCAACATCAAGTTCG
GACCCATTGGCAGCACCGGCAACCCTAGCGGCGGCAACCCTCCCGGCGGAAA
CCCGCCTGGCACCACCACCACCCGCCGCCCAGCCACTACCACTGGAAGCTCT
CCCGGACCTACTAGTAAGCGGATAAGGCGCGCCGCGCGCCAGCTCCGTGCGA
AAGCCTGACGCACCGGTAGATTCTTGGTGAGCCCGTATCATGACGGCGGCGG
GAGCTACATGGCCCCGGGTGATTTATTTTTTTTGTATCTACTTCTGACCCTTTT
CAAATATACGGTCAACTCATCTTTCACTGGAGATGCGGCCTGCTTGGTATTGC
GATGTTGTCAGCTTGGCAAATTGTGGCTTTCGAAAACACAAAACGATTCCTTA
GTAGCCATGCATTTTAAGATAACGGAATAGAAGAAAGAGGAAATTAAAAAA
AAAAAAAAAACAAACATCCCGTTCATAACCCGTAGAATCGCCGCTCTTCGTG
TATCCCAGTACCAGTTTATTTTGAATAGCTCGCCCGCTGGAGAGCATCCTGAA
TGCAAGTAACAACCGTAGAGGCTGACACGGCAGGTGTTGCTAGGGAGCGTCG
TGTTCTACAAGGCCAGACGTCTTCGCGGTTGATATATATGTATGTTTGACTGC
AGGCTGCTCAGCGACGACAGTCAAGTTCGCCCTCGCTGCTTGTGCAATAATC
GCAGTGGGGAAGCCACACCGTGACTCCCATCTTTCAGTAAAGCTCTGTTGGT
GTTTATCAGCAATACACGTAATTTAAACTCGTTAGCATGGGGCTGATAGCTTA
ATTACCGTTTACCAGTGCCGCGGTTCTGCAGCTTTCCTTGGCCCGTAAAATTC
GGCGAAGCCAGCCAATCACCAGCTAGGCACCAGCTAAACCCTATAATTAGTC

TCTTATCAACACCATCCGCTCCCCCGGGATCAATGAGGAGAATGAGGGGGAT

GCGGGGCTAAAGAAGCCTACATAACCCTCATGCCAACTCCCAGTTTACACTC

GTCGAGCCAACATCCTGACTATAAGCTAACACAGAATGCCTCAATCCTGGGA

AGAACTGGCCGCTGATAAGCGCGCCCGCCTCGCAAAACCATCCCTGATGAA

TGGAAAGTCCAGACGCTGCCTGCGGAAGACAGCGTTATTGATTTCCCAAAGA

AATCGGGGATCCTTTCAGAGGCCGAACTGAAGATCACAGAGGCCTCCGCTGC

AGATCTTGTGTCCAAGCTGGCGGCCGGAGAGTTGACCTCGGTGGAAGTTACG

CTAGCATTCTGTAAACGGGCAGCAATCGCCCAGCAGTTAGTAGGGTCCCCTC

TACCTCTCAGGGAGATGTAACAACGCCACCTTATGGGACTATCAAGCTGACG

CTGGCTTCTGTGCAGACAAACTGCGCCCACGAGTTCTTCCTGACGCCGCTCT

CGCGCAGGCAAGGGAACTCGATGAATACTACGCAAAGCACAAGAGACCCGT

TGGTCCACTCCATGGCCTCCCCATCTCTCTCAAAGACCAGCTTCGAGTCAAGG

TACACCGTTGCCCCTAAGTCGTTAGATGTCCCTTTTGTCAGCTAACATATGC

CACCAGGGCTACGAAACATCAATGGGCTACATCTCATGGCTAAACAAGTACG

ACGAAGGGGACTCGGTTCTGACAACCATGCTCCGCAAAGCCGGTGCCGTCTT

CTACGTCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCAAC

AACATCATCGGGCGCACCGTCAACCCACGCAACAAGAACTGGTCGTGCGGCG

GCAGTTCTGGTGGTGAGGGTGCGATCGTTGGGATTCGTGGTGGCGTCATCGG

TGTAGGAACGGATATCGGTGGCTCGATTCGAGTGCCGGCCGCGTTCAACTTC

CTGTACGGTCTAAGGCCGAGTCATGGGCGGCTGCCGTATGCAAAGATGGCGA

ACAGCATGGAGGGTCAGGAGACGGTGCACAGCGTTGTCGGGCCGATTACGCA

CTCTGTTGAGGGTGAGTCCTTCGCCTCTTCCTTCTTTTCCTGCTCTATACCAGG

CCTCCACTGTCCTCCTTTCTTGCTTTTTATACTATATACGAGACCGGCAGTCAC

TGATGAAGTATGTTAGACCTCCGCCTCTTCACCAAATCCGTCCTCGGTCAGGA

GCCATGGAAATACGACTCCAAGGTCATCCCCATGCCCTGGCGCCAGTCCGAG

TCGGACATTATTGCCTCCAAGATCAAGAACGGCGGGCTCAATATCGGCTACT

ACAACTTCGACGGCAATGTCCTTCCACACCCTCCTATCCTGCGCGGCGTGGAA

ACCACCGTCGCCGCACTCGCCAAAGCCGGTCACACCGTGACCCCGTGGACGC

CATACAAGCACGATTTCGGCCACGATCTCATCTCCCATATCTACGCGGCTGAC

GGCAGCGCCGACGTAATGCGCGATATCAGTGCATCCGGCGAGCCGGCGATTC

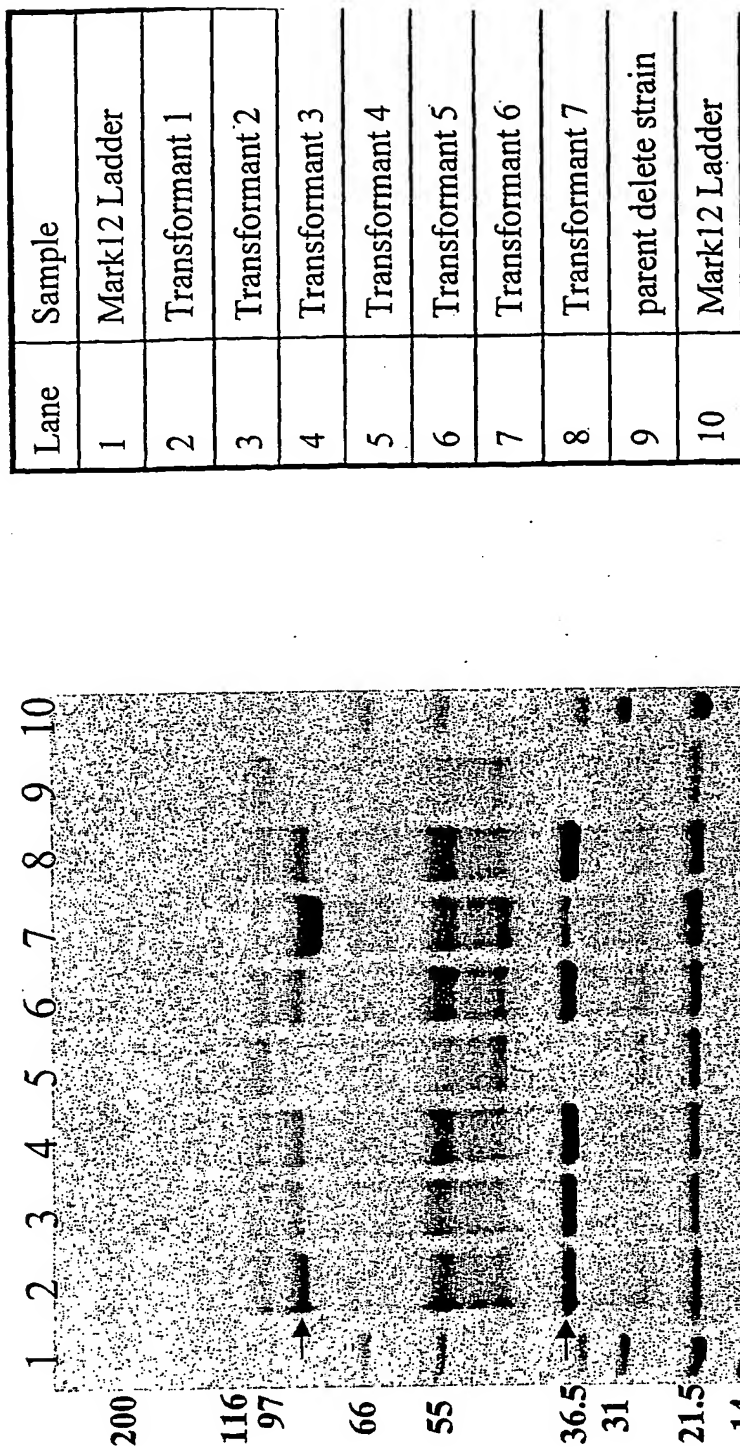CAAATATCAAAGACCTACTGAACCCGAACATCAAAGCTGTTAACATGAACGA

GCTCTGGGACACGCATCTCCAGAAGTGGAATTACCAGATGGAGTACCTTGAG

AAATGGCGGGAGGCTGAAGAAAAGGCCGGGAAGGAACTGGACGCCATCATC

GCGCCGATTACGCCTACCGCTGCGGTACGGCATGACCAGTTCCGGTACTATG

GGTATGCCTCTGTGATCAACCTGCTGGATTTCACGAGCGTGGTTGTTCCGGTT

ACCTTTGCGGATAAGAACATCGATAAGAAGAATGAGAGTTTCAAGGCGGTTA

GTGAGCTTGATGCCCTCGTGCAGGAAGAGTATGATCCGGAGGCGTACCATGG

GGCACCGGTTGCAGTGCAGGTTATCGGACGGAGACTCAGTGAAGAGAGGAC

GTTGGCGATTGCAGAGGAAGTGGGGAAGTTGCTGGGAAATGTGGTGACTCCA

TAGCTAATAAGTGTCAGATAGCAATTTGCACAAGAAATCAATACCAGCAACT

GTAAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAGAGCAGAAAAAAA

CCTGCCGTAGAACCGAAGAGATATGACACGCTTCCATCTCTCAAAGGAAGAA

TCCCTTCAGGGTTGCGTTTCCAGTCTAGACACGTATAACGGCACAAGTGTCTC

TCACCAAATGGGTTATATCTCAAATGTGATCTAAGGATGGAAAGCCCAGAAT

CTAGGCCTATTAATATTCCGGAGTATACGTAGCCGGCTAACGTTAACAACCG

GTACCTCTAGAACTATAGCTAGCATGCGCAAATTTAAAGCGCTGATATCGAT

CGCGCGCAGATCCATATATAGGGCCCGGGTTATAATTACCTCAGGTCGACGT

CCCATGGCCATTCGAATTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT

TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA

AAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTC

ACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCG

GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTC

GCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC

ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAA

AGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCG

CGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAAT

CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG

GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCT

TACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA

GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC

TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA

TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC

ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT

TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTG

CGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG

GCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATT

ACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGT

CTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATT

ATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT

CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATC

AGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTG

ACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCC

AGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAG

CAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTT

ATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTT

CGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTG

TCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG

GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTC

CTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATG

GCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGT

GACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCG

AGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAA

CTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAG

GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACT

GATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA

AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATA

CTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTC

ATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTC

CGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTAT

CATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGC

GTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGT

CACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCG

TCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGC

AGATTGTACTGAGAGTGCACCATAAAATTGTAAACGTTAATATTTTGTTAAAA
TTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAAT
CGGCAAAATCCCTTATAAATCAAAAGAATAGCCCGAGATAGGGTTGAGTGTT
GTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCA
AAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC
CAAATCAAGTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCT
AAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCG
AGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAG
TGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCG
CTACAGGGCGCGTACTATGGTTGCTTTGACGTATGCGGTGTGAAATACCGCA
CAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCTG
CGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGC
TGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTT
TTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGCC

Figure 18

SDS-PAGE gel of supernate samples of shake flask grown *T.reesei* delete strain transformed with the fusion expression construct *cbh1-E1*

Lane	Sample
1	Mark12 Ladder
2	Transformant 1
3	Transformant 2
4	Transformant 3
5	Transformant 4
6	Transformant 5
7	Transformant 6
8	Transformant 7
9	parent delete strain
10	Mark12 Ladder

The fusion protein is indicated by the upper arrow, the cleaved E1 catalytic domain is indicated by the lower arrow
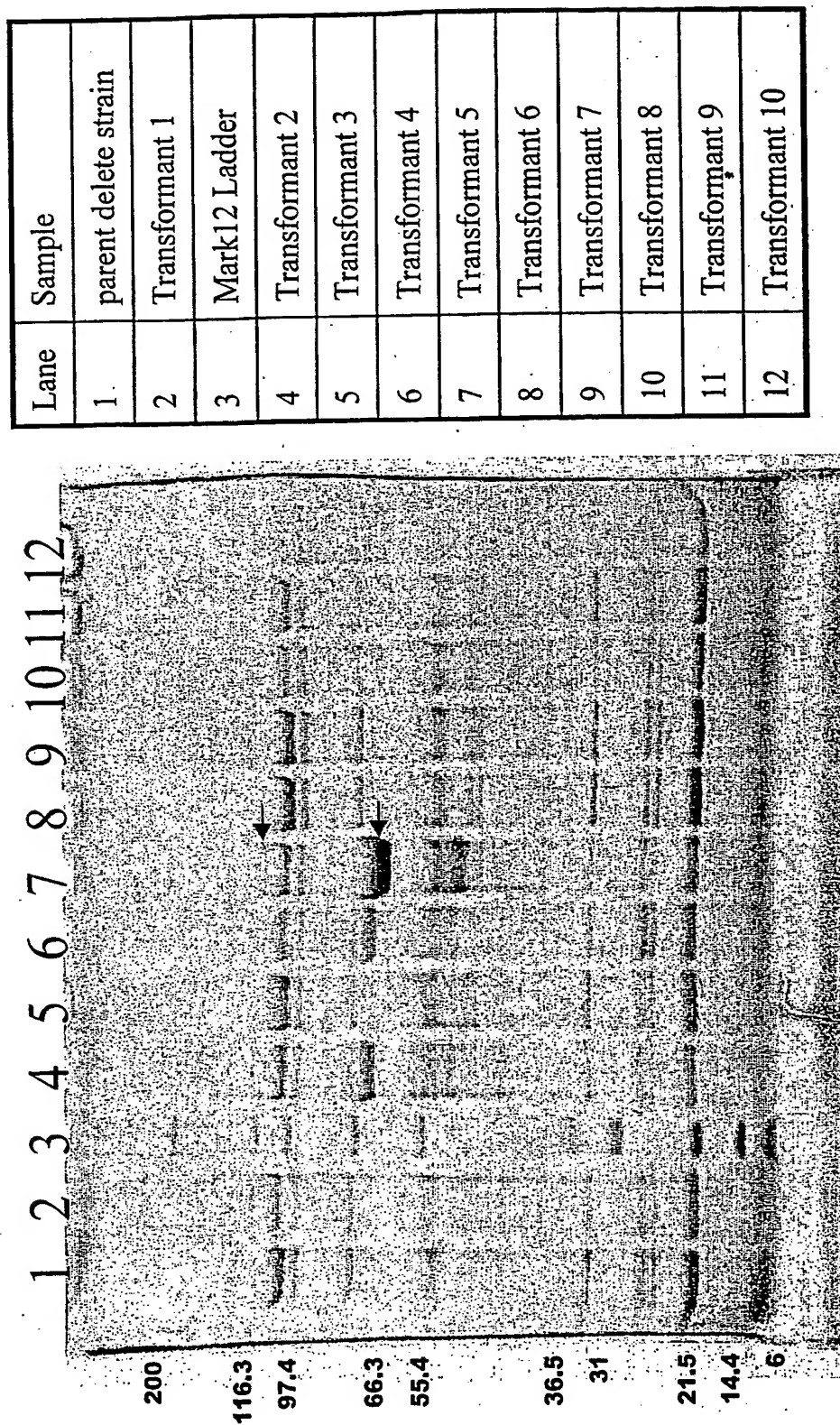
Figure 19

SDS-PAGE gel of supernate samples of shake flask grown *T.reesei* delete strain transformed with the fusion expression construct *cbh1-GH48*
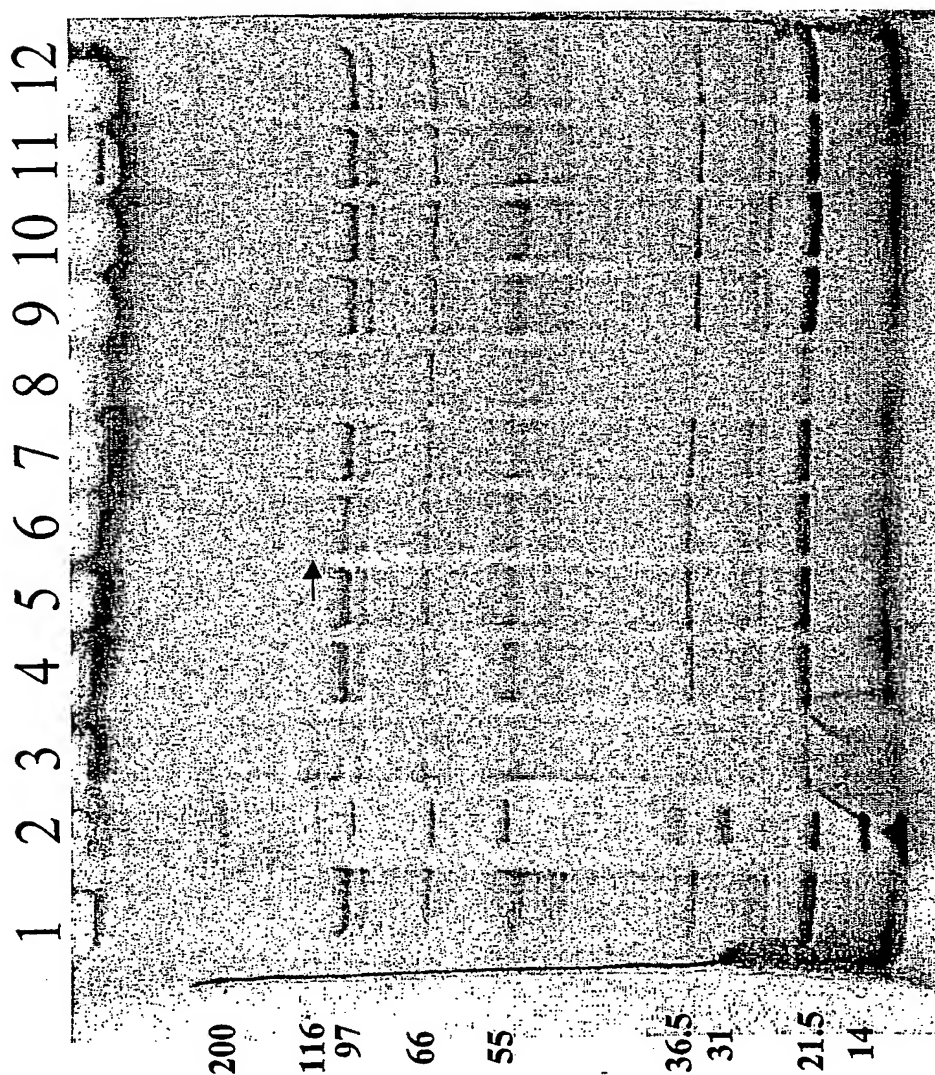
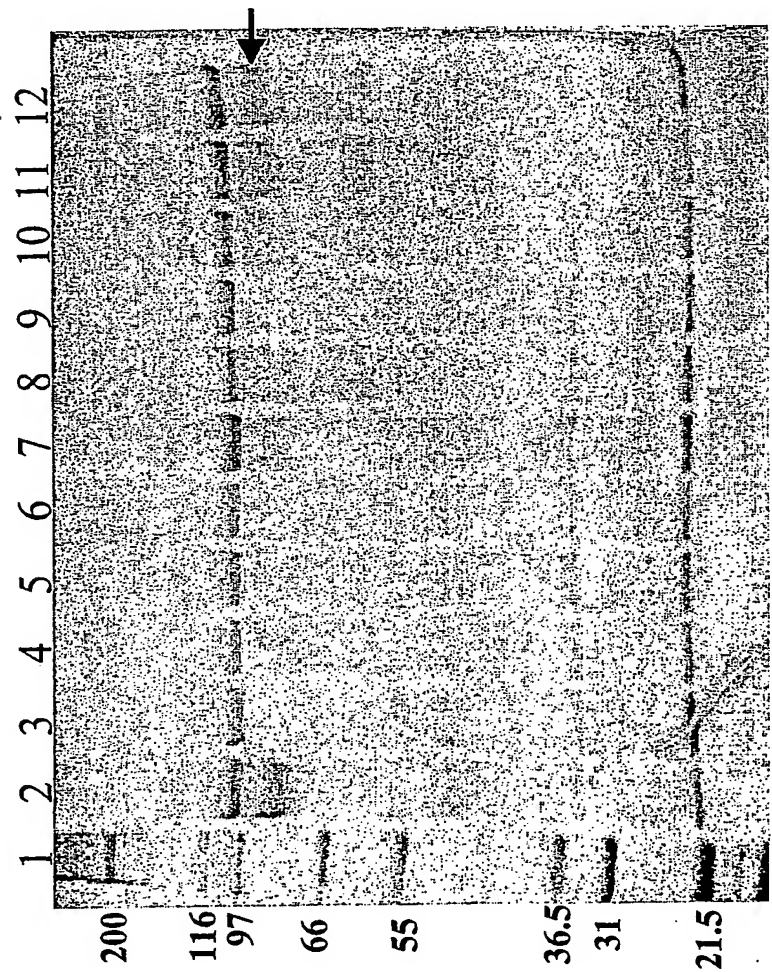| Lane	Sample
1	parent delete strain
2	Mark12 Ladder
3	Transformant 1
4	Transformant 2
5	Transformant 3
6	Transformant 4
7	Transformant 5
8	Transformant 6
9	Transformant 7
10	Transformant 8
11	Transformant 9
12	Transformant 10

The fusion protein is indicated by the arrow

Figure 20

SDS-PAGE gel of supernate samples of shake flask grown *T.reesei* delete strain
transformed with the fusion expression construct *cbh1-GH74*

Lane	Sample
1	parent delete strain
2	Transformant 1
3	Mark12 Ladder
4	Transformant 2
5	Transformant 3
6	Transformant 4
7	Transformant 5
8	Transformant 6
9	Transformant 7
10	Transformant 8
11	Transformant 9
12	Transformant 10

The fusion protein is indicated by the upper arrow, the cleaved GH74 catalytic domain is indicated
by the lower arrow

Figure 21

SDS-PAGE gel of supernate samples of shake flask grown *T.reesei* delete strain
transformed with the fusion expression construct *cbh1-E3*

Lane	Sample
1	Mark12 Ladder
2	Transformant 1
3	Transformant 2
4	Transformant 3
5	Transformant 4
6	Transformant 5
7	Transformant 6
8	Transformant 7
9	Transformant 8
10	Transformant·9
11	Transformant 10
12	Transformant 11

arrow indicates new protein expressed
in cbh1-E3 transformants

Figure 22

SDS-PAGE gel of supernate samples of shake flask grown *T.reesei* delete strain transformed with the fusion expression construct cbh1-E5

Lane	Sample
1	Mark12 Ladder
2	parent delete strain
3	Transformant 1
4	Transformant 2
5	Transformant 3
6	Transformant 4
7	Transformant 5
8	Transformant 6
9	Transformant 7
10	Transformant 8
11	Transformant 9
12	Transformant 10
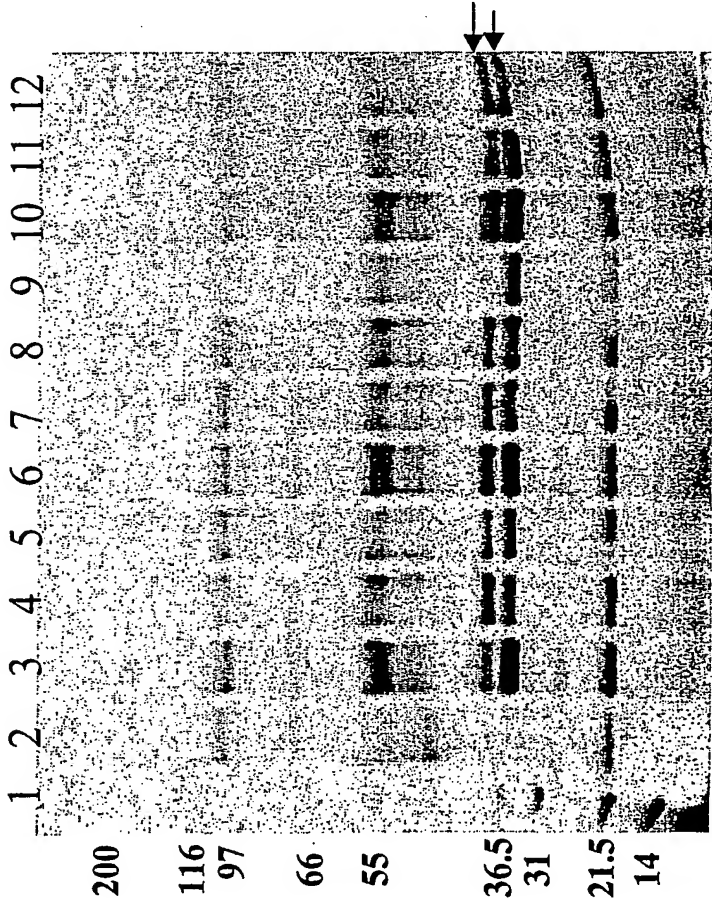
E5 is indicated by the lower arrows

FIGURE 23
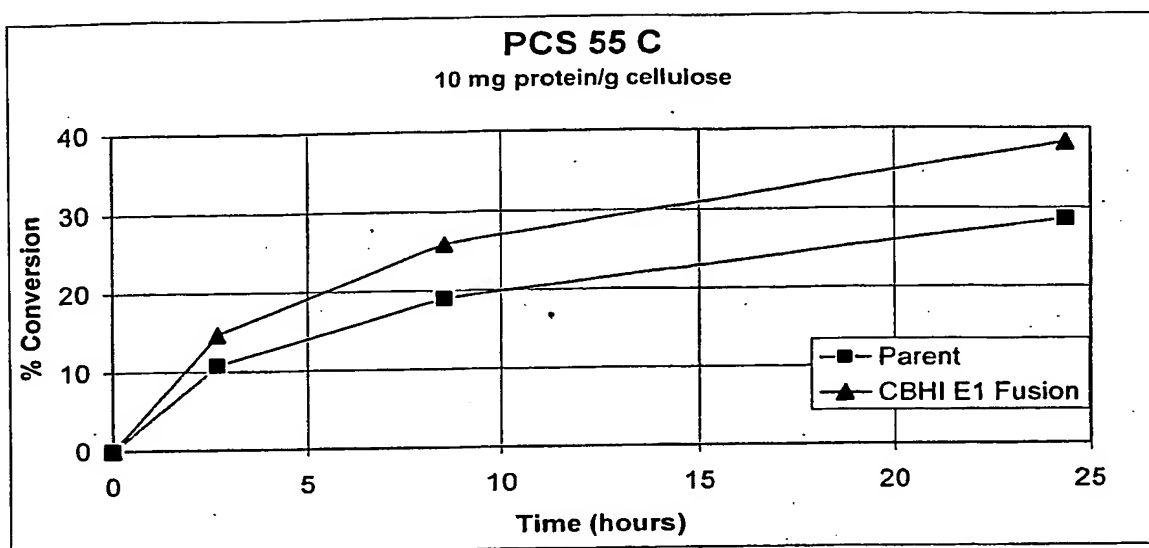
This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

❏ **BLACK BORDERS**

❏ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

❏ **FADED TEXT OR DRAWING**

☑ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

❏ **SKEWED/SLANTED IMAGES**

❏ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

❏ **GRAY SCALE DOCUMENTS**

❏ **LINES OR MARKS ON ORIGINAL DOCUMENT**

❏ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

❏ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.